B1 building the decision tree based on the corresponding set of one or more fuzzy clusters.

REMARKS

By this amendment, claims 1-34 are pending, in which claims 1, 9, 18, and 26 are amended. Care was exercised to avoid the introduction of new matter.

The Office Action mailed August 28, 2002 rejected claims 1-5, 9-13, 16-22, 26-30, and 33-24 obvious under 35 U.S.C. § 103 based on *Hall et al.* (Hall et al., "Generating Fuzzy Rules from Data," *IEEE*, 1996), claims 6, 14, 23, 31 as obvious over *Hall et al.* in view of *Shafer et al.* (Shafer et al., "SPRINT: A Scalable Parallel Classifier for Data Mining," *Proceedings of the 22nd VLDB Conference*, 1996), and claims 7-8, 15, 24-25, and 32 as obvious over *Hall et al.* in view of *Choe et al.* (Choe et al., "On the Optimal Choice of Parameters in a Fuzzy C-Means Algorithm," *IEEE*, 1992). These rejections are respectfully for at least the following reasons.

CLAIMS 1-9 AND 18-26

The rejection of claims 1-9 and 18-26 is respectfully traversed because *Hall et al.*, individually or in combination with *Shafer et al.* and *Choe et al.*, fail to teach or otherwise the limitations of the claims. For example, method claim 1 (whose limitations are mirrored in computer-readable medium claim 18) sets forth:

> 1. (Once Amended) A method for refining a node of a decision tree associated with a plurality of data characterized by a plurality of features, comprising:
> selecting a feature from among the features characterizing the data associated with the node;
> performing a cluster analysis along the selected feature to group the data into one or more clusters; and
> constructing one or more arcs of the decision tree at the node respectively for each of the one or more clusters.

Accordingly, claims 1 and 18 recites a way of refining a node in a decision tree by selecting a feature of those that characterize the data associated with the node, then performing a

cluster analysis along the selected feature, and then constructing arcs of the decision tree for each of the clusters. Thus, a cluster analysis is performed in refining a node in a decision tree, enabling the decision to be built "on the fly" (see Spec., p. 6).

By contrast, *Hall et al.* does not show this way of building a decision tree, by performing a cluster analysis in refining a node of a decision tree. Rather, *Hall et al.* is directed to a method of developing of fuzzy rules from continuous valued data by building a decision tree in accordance with the C4.5 algorithm (Abstract, p. 1757, col. 1). However, *Hall et al.* recognize that the "C4.5 algorithm tree algorithm **requires crisp** class assignments for all objects. It is **necessary** to partition the continuous output values into a effect set of **discrete** output classes." (Section 2.1, p. 1758, col. 1, emphasis added). Accordingly, *Hall et al.* propose to preprocess the data first by applying a fuzzy c-means clustering to determine the discrete classes, and then feeding the discrete classes into the C4.5 algorithm: "After a discrete class has been created for each example, as discussed in Section 2.1, C4.5 may be used to create a decision tree." (Section 3, p. 1759, col. 1).

Accordingly, *Hall et al.* fails to teach or suggest "performing a cluster analysis along the **selected** feature to group the data into one or more clusters" since whatever cluster analysis that is performed in *Hall et al.* is performed before building the decision tree, that is, without selecting a feature when refining a node of a decision tree. The remaining references, *Shafer et al.* and *Choe et al.*, also fail to teach this aspect of claims 1-9 and 18-26.

Dependent claims 2-9 and 19-26 include these limitations by their dependency are therefore allowable for at least the same reasons as their independent claims 1 and 18, respectively. Moreover, dependent claims 9 and 26 provide that the steps of selecting the feature and performing the cluster analysis is performed recursively, which is not taught in *Hall et al.*

since *Hall et al.*'s clustering is performed before invoking the C4.5 algorithm, recursively within the C4.5 algorithm.

CLAIMS 2-3, 10-17, 19-20, AND 27-34

Hall et al., alone or in combination with *Shafer et al.* and *Choe et al.*, fail to teach or suggest the limitations of claims 2-3, 10-17, 19-20, and 27-34. For example, independent claims 10 and 27 recite:

> performing a plurality of cluster analyses along each of the features to calculate a **maximal cluster validity measure**, said maximal cluster validity measure corresponding to one of the features;
> selecting the one of the features corresponding to the maximal cluster validity measure;

Dependent claims 2 and 19 also affirmatively recite these limitations. None of the references show the recited "maximal cluster validity measure" calculated by performed a plurality of cluster analyses and selecting one of the features that corresponds to the maximal cluster validity measure. Moreover, claims 3, 11, 17, 20, 27, and 34 specify a specific kind of maximal cluster validity measure, that based on the "partition coefficient."

As explained above, *Hall et al.* discloses a method of generating fuzzy rules from data by first performing a fuzzy cluster analysis to determine crisp, discrete classes for the data and then applying the C4.5 decision tree algorithm to the discrete classes. Since the C4.5 decision tree algorithm requires discrete classes, the C4.5 algorithms selects its features to build the decision tree based on the "highest information gain associated with it" (Section 2, p. 1758, col. 1)–but not on a "maximal cluster validity measure" or a "partition coefficient" based on performing cluster analyses as recited in the claims.

The remaining references, *Shafer et al.* and *Choe et al.*, also fail to teach this aspect of claims 2-3, 10-17, 19-20, and 27-34 and were not cited for that purpose.

CLAIMS 7-8, 15, 24-25, AND 32

The dependent claims are allowable for at least the same reasons as their independent claims and are individually on their own merits. For example, dependent claims 7-8, 15, 24-25, and 32 cover the element of "calculating a domain **ratio** of a difference in domains limits of the data **over** a difference in domain limits of a superset of the data." Accordingly, the domain ratio is recited to be a ratio of one difference over another.

The Office Action recognized correctly that *Hall et al.* fails to disclose this element, but incorrectly relies on *Choe et al.* for this feature. Specifically, the Office Action cites step 6 of *Choe et al.*'s algorithm, which states: "return to Step 3 if $\| U^{(l+1)} - U^{(l)} \| > \varepsilon$."[1] However, this condition is a **difference** of two quantities, $U^{(l+1)}$ and $U^{(l)}$, not a ratio of two differences as recited in claims 7-8, 15, 24-25, and 32.

Therefore, the present application, as amended, overcomes the rejections of record and is in condition for allowance. Favorable consideration is respectfully requested. If any unresolved issues remain, it is respectfully requested that the Examiner telephone the undersigned attorney at 703-425-8516 so that such issues may be resolved as expeditiously as possible.

Respectfully Submitted,

DITTHAVONG & CARLSON, P.C.

1/28/2003
Date

Stephen C. Carlson
Attorney/Agent for Applicant(s)
Reg. No. 39929

10507 Braddock Rd, Suite A
Fairfax, VA 22032
Tel. 703-425-8516
Fax. 703-425-8518

[1] Due to typographical limitations, the "not less than or equal to" symbol ($\not\leq$ with a stroke through it) of *Choe et al.* is replaced by the equivalent "greater than" (>) symbol.

APPENDIX

1. (Once Amended) A method for [generating] refining a node of a decision tree [for] associated with a plurality of data characterized by a plurality of features, comprising:

selecting a feature from among the features characterizing the data associated with the node;

performing a cluster analysis along the selected feature to group the data into one or more clusters; and

[building the decision tree based on] constructing one or more arcs of the decision tree at the node respectively for each of the one or more clusters.

9. (Once Amended) The method according to claim 1, [wherein building the decision tree based on the one or more clusters includes] further comprising the steps of:

projecting the data in each of the clusters, wherein the projected data are characterized by the plurality of the features but for the selected feature; and

recursively performing the steps of selecting a feature and performing the cluster analysis on the projected data in each of the clusters.

18. (Once Amended) A computer-readable medium bearing instructions for [generating] refining a node of a decision tree [for] associated with a plurality of data characterized by a plurality of features, said instructions being arranged to cause one or more processors upon execution thereby to perform the steps of:

selecting a feature from among the features characterizing the data associated with the node;

performing a cluster analysis along the selected feature to group the data into one or more clusters; and

[building the decision tree based on] constructing one or more arcs of the decision tree at the node respectively for each of the one or more clusters.

26. (Once Amended) The computer-readable medium according to claim 18, wherein [building the decision tree based on the one or more clusters includes] said are further arranged to the one or more processors upon execution thereby to perform the steps of:

projecting the data in each of the clusters, wherein the projected data are characterized by the plurality of the features but for the selected feature; and

recursively performing the steps of selecting a feature and performing the cluster analysis on the projected data in each of the clusters.